

As a library, NLM provides access to scientific literature. Inclusion in an NLM database does not imply endorsement of, or agreement with, the contents by NLM or the National Institutes of Health.

Learn more: [PMC Disclaimer](#) | [PMC Copyright Notice](#)



Microbiome. 2022 Aug 24;10:134. doi: [10.1186/s40168-022-01332-w](https://doi.org/10.1186/s40168-022-01332-w)

Machine learning algorithm to characterize antimicrobial resistance associated with the International Space Station surface microbiome

[Pedro Madrigal](#)^{1,2,✉}, [Nitin K Singh](#)³, [Jason M Wood](#)³, [Elena Gaudioso](#)⁴, [Félix Hernández-del-Olmo](#)⁴, [Christopher E Mason](#)^{5,6,7,8}, [Kasthuri Venkateswaran](#)^{3,#}, [Afshin Beheshti](#)^{9,10,#}

[Author information](#) [Article notes](#) [Copyright and License information](#)

PMCID: PMC9400218 PMID: [35999570](#)

Abstract

Background

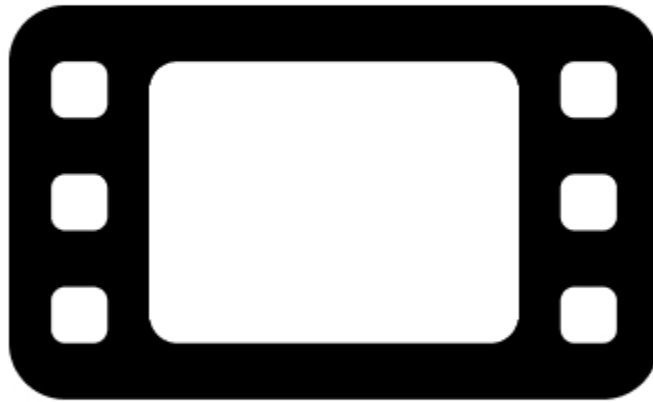
Antimicrobial resistance (AMR) has a detrimental impact on human health on Earth and it is equally concerning in other environments such as space habitat due to microgravity, radiation and confinement, especially for long-distance space travel. The International Space Station (ISS) is ideal for investigating microbial diversity and virulence associated with spaceflight. The shotgun metagenomics data of the ISS generated during the Microbial Tracking–1 (MT-1) project and resulting metagenome-assembled genomes (MAGs) across three flights in eight different locations during 12 months were used in this study. The objective of this study was to identify the AMR genes associated with whole genomes of 226 cultivable strains, 21 shotgun metagenome sequences, and 24 MAGs retrieved from the ISS environmental samples that were treated with propidium monoazide (PMA; viable microbes).

Results

We have analyzed the data using a deep learning model, allowing us to go beyond traditional cut-offs based only on high DNA sequence similarity and extending the catalog of AMR genes. Our results in PMA treated samples revealed AMR dominance in the last flight for *Kalamiella piersonii*, a bacteria related to urinary tract infection in humans. The analysis of 226 pure strains isolated from the MT-1 project revealed hundreds of antibiotic resistance genes from many isolates, including two top-ranking species that corresponded to strains of *Enterobacter bugandensis* and *Bacillus cereus*. Computational predictions were experimentally validated by antibiotic resistance profiles in these two species, showing a high degree of concordance. Specifically, disc assay data confirmed the high resistance of these two pathogens to various beta-lactam antibiotics.

Conclusion

Overall, our computational predictions and validation analyses demonstrate the advantages of machine learning to uncover concealed AMR determinants in metagenomics datasets, expanding the understanding of the ISS environmental microbiomes and their pathogenic potential in humans.



0:00 / 0:00



[Download video file](#) (99.9MB, mp4)

[Open in a new tab](#)

Video Abstract

Supplementary Information

The online version contains supplementary material available at 10.1186/s40168-022-01332-w.

Keywords: ISS, Metagenomics, Antibiotic resistance, Machine learning, Space Omics, Microbiome, Built-environment, Microbial Tracking-1, NGS

Background

According to the World Health Organization, the widespread use of antibiotics worldwide and the slow discovery of major types of antibiotics in the last thirty years has made antibiotic resistance one of the biggest threats to human health, food security, and development [58]. Accordingly, with NASA setting the course to return to the Moon with the Artemis mission and eventually venture out to Mars, maintaining the health of astronauts during long-term spaceflight is of paramount importance [1]. One area of particular concern is the reported increase in virulence and antibiotic resistance of microorganisms in space experiments [4, 30, 36, 49, 51, 56, 61]. Combined with a depressed or altered immune response in astronauts [25, 46], there is an increased risk of opportunistic microbial infection. Spaceflight promotes biofilm formation [32], and bacteria cultured from astronauts during flight were more resistant than isolates obtained from the same individual either pre- or post-flight [50]. Mutations also occurred more frequently in long-term spaceflights [24]. An alternative non-mutually exclusive hypothesis to increased virulence or microbial resistance to antibiotics is that spaceflight conditions might alter the stability of pharmaceuticals [23]. In any case, bacterial infections might be more challenging to treat in space.

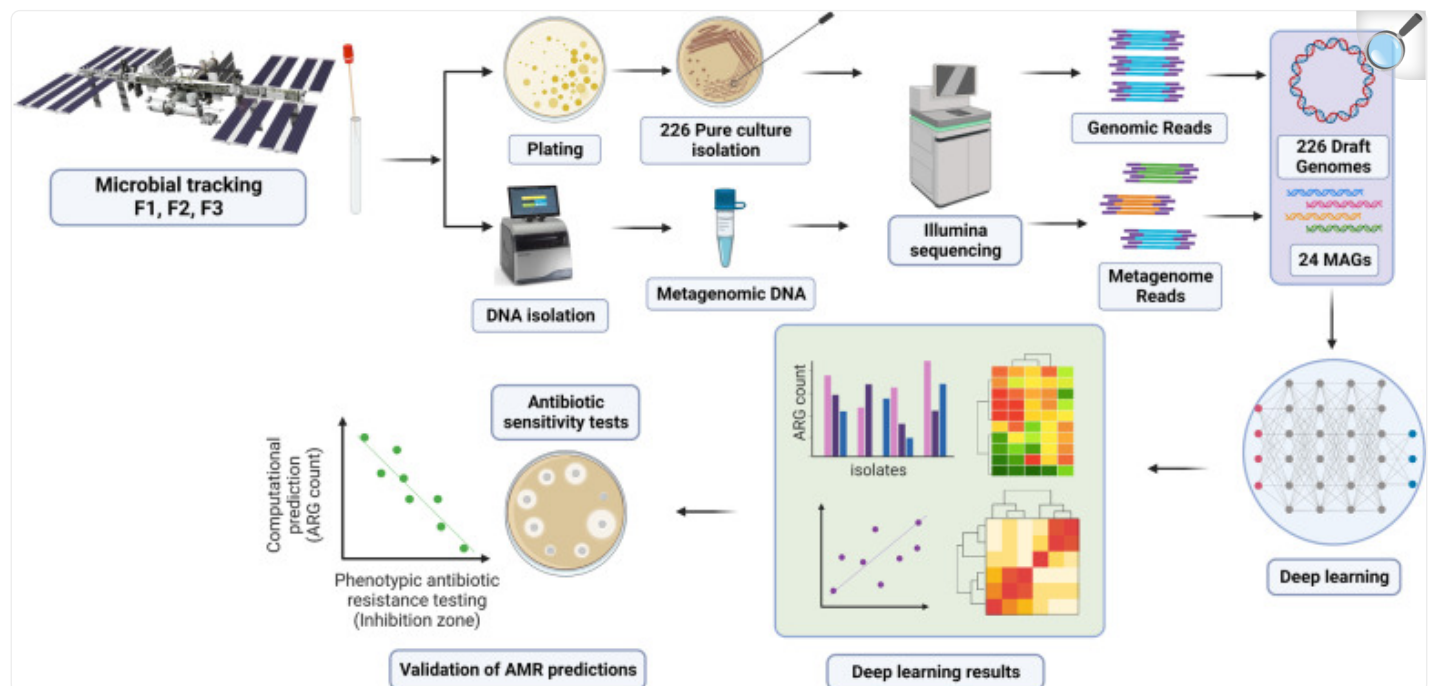
The International Space Station (ISS) is a closed-built environment with its own environmental microbiome shaped by microgravity, radiation, and limited human presence [53]. We and others have shown that microbiomes are dynamic, diverse and sometimes intertwined at the ISS. Be et al. [8] analyzed antibiotic resistance and virulence genes from dust and vacuum filter samples of ISS (treated with propidium monoazide, or PMA), demonstrating that human skin-associated microbes impact the ISS microbiome. Indeed, the skin and intestinal microbiomes of astronauts that spent 6 to 12 months in the ISS have been shown to be altered [55]. In addition, the salivary microbiome of astronauts changed as a result of spaceflight, potentially activating microbes that promote viral replication [52] and altering the abundance of some antimicrobial resistance (AMR) genes [35]. The ISS itself also presents specific core microbiome signatures on its surfaces that we characterized recently using shotgun metagenome and amplicon sequencing [14, 43, 51], analogous to microbiome signatures found in specific geographies on Earth [18].

Further analyses across several missions have revealed that the microbiome of the crew's skin resembled those of the surfaces inside the ISS collected by the crewmember on the same flight [5]. To better understand the composition of these bacterial populations we and others have characterized shotgun whole-genome sequencing (WGS) of several ISS microorganisms [10, 11, 45]. Although most of them have been found to be non-pathogenic to humans, there are exceptions such as antibiotic-resistant *Enterobacter bugandensis* strains that could have an increased chance of pathogenicity [44].

Computational analyses of microbiome data collected in Earth have shown that AMR can be predicted from genomic sequence of pure cultures alone [29, 47], but a consensus approach on the best way to detect AMRs in metagenomic datasets has yet to be established [41]. Generally, predictions are restricted to high identity (high sequence similarity to databases) cut-offs, requiring a 'best-hit' on an appropriate AMR database with a sequence identity greater than 80% by many programs such as ResFinder [60]. Although the 'best-hit' approach has a low false-positive rate, the false-negative rate can be very high, and a large number of actual Antibiotic Resistance Genes (ARGs) are predicted as non-ARGs,

thus concealing the identification of potentially functional ARGs [3]. Another method of identification is to link the immune repertoire of the astronaut to the peptides of the microbes on the ISS, but this requires complex coordination with crew sampling and is rare [19]. However, it has been shown recently that deep learning, a class of machine learning algorithms, can expand the catalog of AMR genes and increase the accuracy of the predictions based on metagenomic data [3, 13, 27]. We then hypothesized that the characterization of AMR from sequencing data at the ISS could be investigated from an artificial intelligence perspective using a robust deep learning framework. For that, we analyzed whole-genome sequences of 226 pure strains (cultivable microbes), metagenome sequences of 21 environmental samples, and 24 MAGs retrieved from PMA treated samples (Fig. 1) using the supervised deep learning approach proposed by Arango-Argoty et al. [3], which has shown high sensitivity for detection of AMR genes in an independent benchmark [57].

Fig. 1.



[Open in a new tab](#)

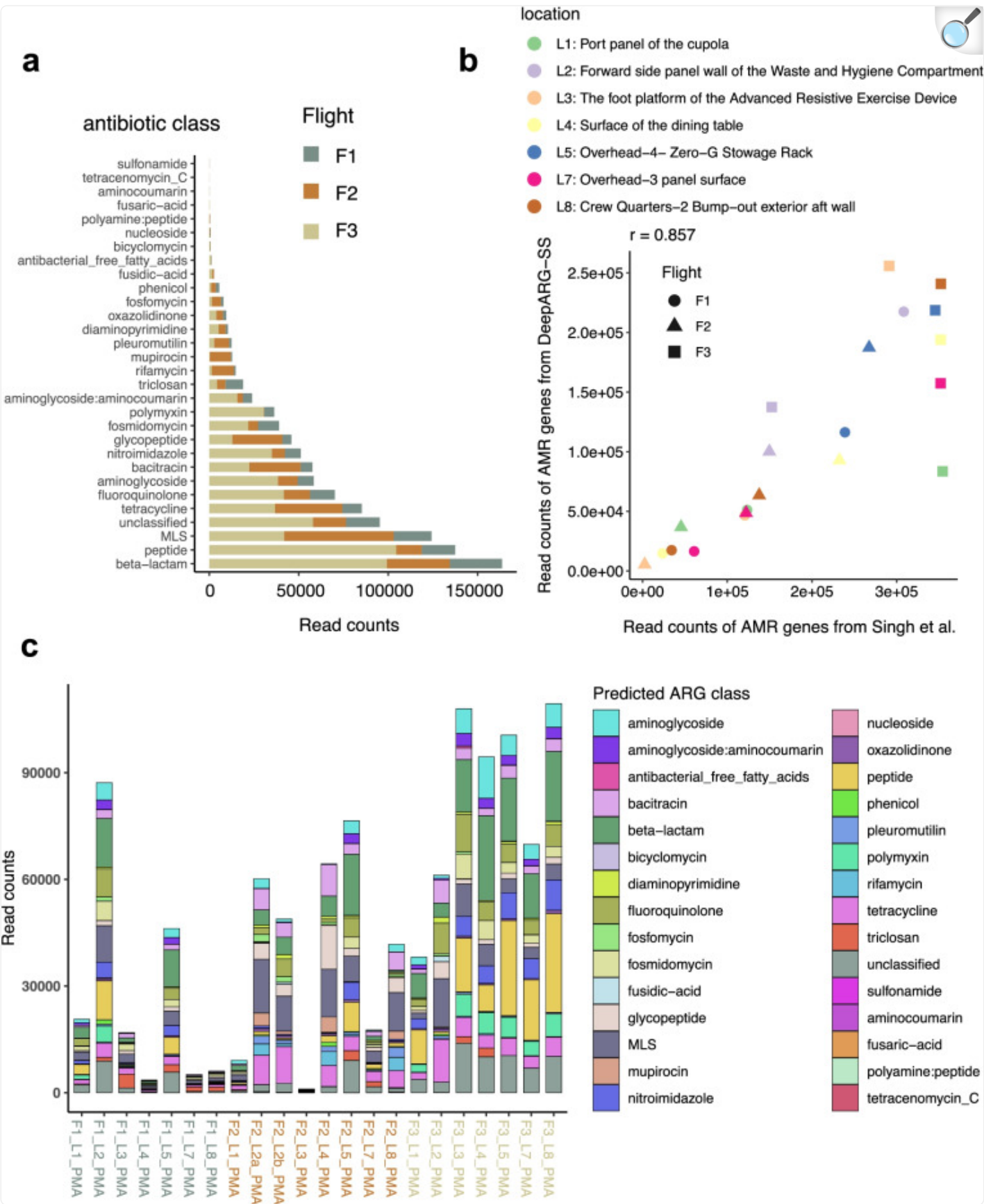
Overview of sample collection and data analysis for the characterization of antibiotic resistance at the ISS using deep learning. The data are processed in a step-wise fashion including data QC, mapping, quantification, and matching to time of collection and mission. The figure has been generated using BioRender (<http://biorender.com>)

Results

Predictions based on short metagenomics sequences and ORFs partly overlap with previous analyses and reveal new AMR determinants at the ISS surface microbiome

The first shotgun metagenome sequencing of intact microbial cells (Propidium monoazide-PMA treated) without whole-genome amplification was performed by Singh et al. [43]. There, samples were taken in 8 locations across three flights (F1, F2, F3) during a period of 12 months. A detailed description of sampling procedures and locations, species diversity and functional characterization can be found in Singh et al. [43]. To deploy a deep learning approach for predicting antibiotic resistance genes from metagenomic data, we used DeepARG, a computational resource proven to be more accurate than traditional approaches [3]. We first run DeepARG-SS (DeepARG for short reads) using the recommended prediction probability cut-off of 0.8 to obtain read counts of AMR genes (Fig. 2a). As in the seminal paper [43], quantification of antibiotics associated with AMR revealed ‘beta lactams’ ranking first and ‘peptide’ second, and generally more AMR reads counts observed in Flight 3 (F3) than in previous two flights (Fig. 2a). However, reads counts in certain antibiotics such as pleuromutilin, mupirocin and rifamycin were found largely in Flight 2 (Fig. 2a). Our read counts correlate ($r = 0.86$, $p = 6.879\text{e-}7$; Pearson’s product-moment correlation) with read counts obtained for antimicrobial resistance by Singh et al. [43] (Fig. 2b). Taken together, these suggest a partial overlap with results obtained in Singh et al. [43] analyzed using the traditional approach.

Fig. 2.

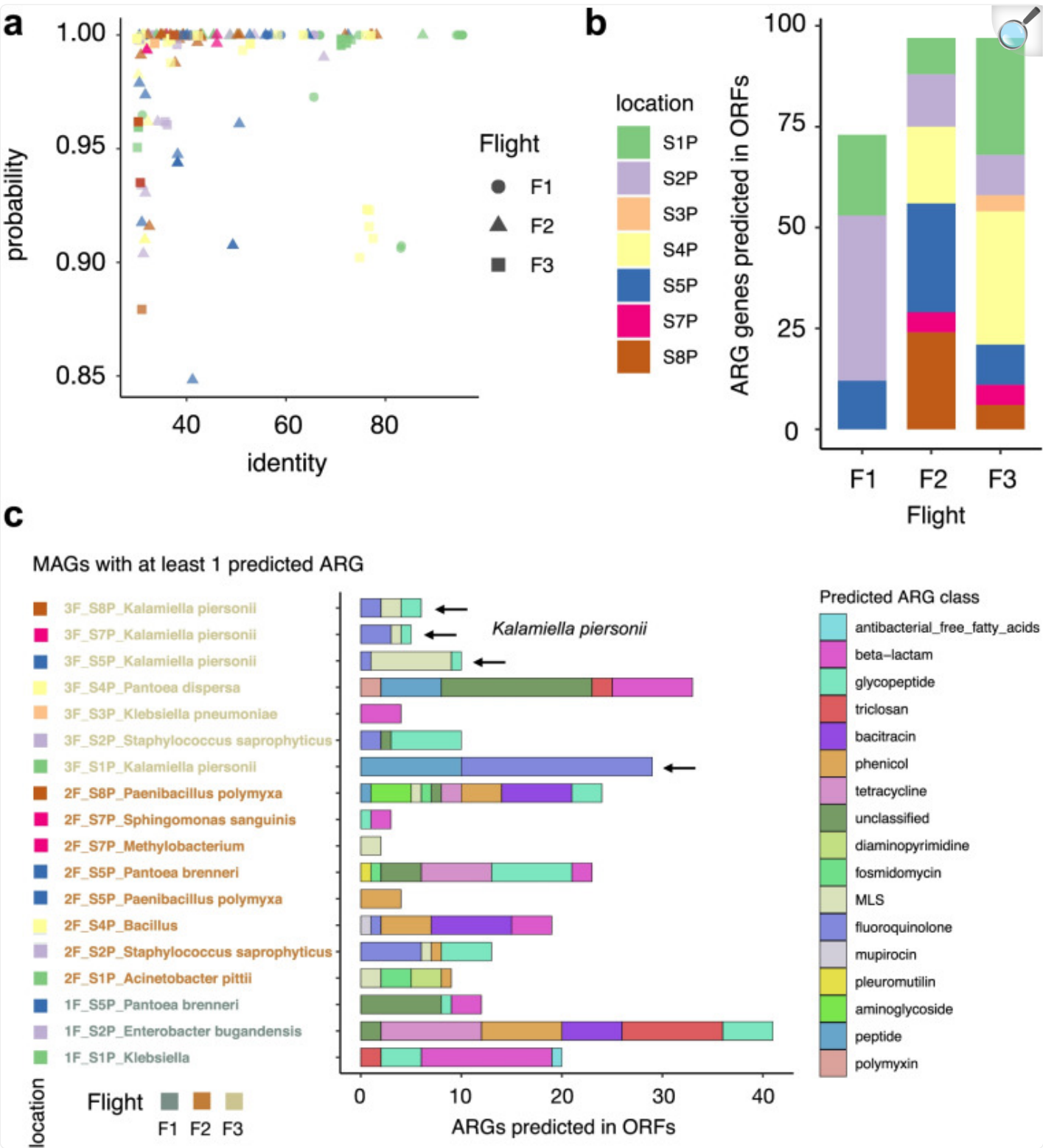


Prediction of ARGs using a pre-trained DeepARG-SS model. **a** Distribution of ARG read counts across antibiotic classes for the three flights (F1, F2, F3). **b** Correlation of read counts found by DeepARG-SS and those in Singh et al. [43]. Pearson's product-moment correlation $r = 0.86$, ($p = 6.879\text{e-}07$) for the three flights and their locations. **c** Read counts of ARG class across flights for each location for PMA-treated samples in Singh et al. [43, 44]. The antibiotic class (*multi-drug*) is not shown. Results are for ARGs with probability > 0.8

While more AMR reads counts were found in Flight 3, we also observed variability between the different locations and flights, and an increasing number of read counts associated with time. For instance, location 4 (L4, surface of the dining table) increased the number of AMR reads counts with successive flights (Fig. 2b, c). While resistance to 'beta lactams' was evenly distributed across flights and locations, resistance to 'polymyxin' and especially 'peptide' represents a more significant proportion of AMR counts in locations of Flight 3 (Fig. 2c). In addition, we also observed the widespread presence of reads related to Macrolides, Lincosamides, Streptogamines (MLS), and tetracycline resistance.

To investigate the possible association between AMR patterns and specific microbes, we assembled the short reads into Metagenome-Assembled Genomes (MAGs; see Methods), identified their Open Reading Frames (ORFs), and repeated the prediction of ARGs using DeepARG-LS [3]. Figure 3a shows the distribution of DeepARG classification probabilities and best-hit identity of ARGs in MAGs from the ISS. As we can retrieve highly probable ARGs (probability > 0.8) presenting low sequence identity (for many ARGs, identity is < 40%), this method is likely more advantageous than using the 'best-hit' approach only. Compared to DeepARG-SS results obtained previously, the analysis of MAGs did not reveal significant differences in the number of ARGs predicted in the ORFs for the different flights (Fig. 3b). However, interestingly the results show a smaller number bacterial species having ARGs in Flight #1 (F1) when compared to Flights #2 and #3 (Fig. 3b, c) (data is shown for MAGs with at least 1 predicted ARG; the total number of MAGs analyzed is 24). Specifically, the number of locations is smaller in Flight 1 ($n = 3$) than in F2 ($n = 6$) and F3 ($n = 7$) (Fig. 3b). Many ARGs were identified in *Kalamiella piersonii* MAGs in multiple locations during F3, showing AMR patterns related to (glyco)peptide, fluoroquinolone and MLS (Fig. 3c). Of note, the *K. piersonii* strain closely related to one found at the ISS has been associated to human urinary tract infection [40]. The potentially very pathogenic microbe *E. bugandensis* was found in location 2 (forward side panel wall of the Waste and Hygiene Compartment) in Flight 1, presenting more than 40 ARGs. In addition, in the original study, *Pantoea* species were found to be the dominant genus in samples in 5 out of 7 locations sampled from Flight 3, especially at location 5 (surface rack). In our re-analysis, we observed *Pantoea brenneri* and *Pantoea dispersa* having ARGs related to beta-lactams and peptide [43], as well as to triclosan and polymyxin resistance.

Fig. 3.



[Open in a new tab](#)

ARGs detected in ORFs in metagenome-assembled genomes (MAGs) from PMA-treated samples. **a**

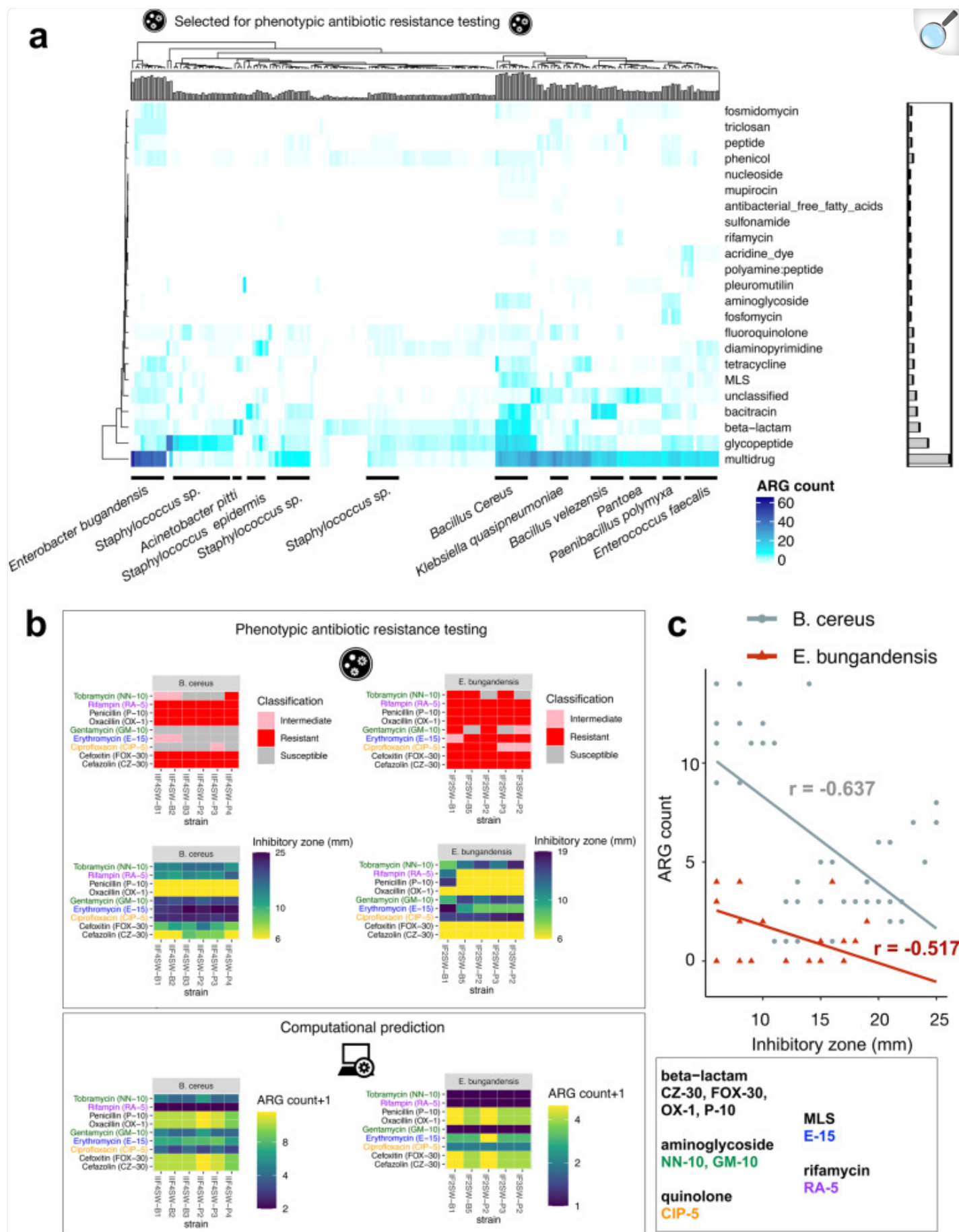
Distribution of DeepARG classification probability and best-hit identity in MAGs retrieved from the ISS. **b** Total number of ARGs predicted for each flight and location. **c** Number of ARGs predicted for each MAG. Most common antibiotic class (*multi-drug*) not shown. The black arrows indicate *Kalamella piersonii*

Overall, our results partially agree with earlier findings while providing new insights into previously unobserved antibiotic resistance classes (of the 30 antibiotic resistance categories included in the model). Specifically, the re-analysis of short sequences and MAGs from the ISS reveals dominance of *K. piersonii* antibiotic resistance in different locations of Flight 3 (Fig. 3c).

Distribution of antibiotic resistance genes in scaffolds of Microbial Tracking-1 strains isolated from the ISS

We then applied DeepARG-LS to 226 Microbial Tracking-1 (MT-1) isolates (Mason and Venkateswaran labs, published and unpublished WGS of MT-1 pure strains isolated from ISS environment). We found a range of 2 to 92 ARGs in 184 out of 226 isolates (Fig. 4a; Table S1). This machine learning approach allowed us to go beyond the traditional cut-off based only on high sequence DNA similarity (Figure S1). These results suggest a widespread presence of potential ARGs in the isolates, with ‘multi-drug’ class being first, followed by glycopeptides, beta-lactams, bacitracin, and tetracyclines. The ‘multi-drug’ antibiotic class was defined by aggregating several antibiotic names from the CARD and ARDB databases (efflux, multi-drug and na_antimicrobials). We then used BLAST to match isolates showing AMR sequences predicted by DeepARG to microbial species (Fig. 4a) and identified *Bacillus cereus* and *E. bugandensis*, which were previously profiled organisms on the ISS [44, 54] as the top 2 ranking species with a high number of ARGs. We have previously shown that five *E. bugandensis* isolates were almost equivalent to nosocomial earth isolates showing resistance to multi-drug antibiotic compounds, fluoroquinolones, and fosfomycin [44]. In addition, *E. bugandensis* strains were shown to be resistant to 9 antibiotics [51]. Our results reinforce the potential pathogenicity of this microbe. Nonetheless, antimicrobial resistance was not examined for *B. cereus* strains in Venkateswaran et al. [54]. *B. cereus* is a food poisoning microorganism that might be a concern for crew members' health. In addition, we found novel ARGs associated with other species such as *K. pneumoniae*, *Pantoea*, *Paenibacillus polymyxa*, *Bacillus velezensis*, *Enterococcus faecalis*, *Sphingomonas*, and, with a lower number of ARGs, several species of *Staphylococcus*. *E. faecalis* virulence was previously shown to be affected by microgravity [28]. We then used the tool Prokka [42] to fully annotate the bacterial isolates, finding as expected that the number of coding sequences, but not the number of ARGs, increased in proportion to genome sizes (Figure S2a). Then, we ran the pan-genome analysis tool Roary [38] to compare isolates of *E. bugandensis* (10) and *B. cereus* (10) finding that the core set of genes was highly conserved among strains of the same species (Figure S2b).

Fig. 4.



Heatmap and clustering of ARG counts detected in MT-1 pure strains isolated from the ISS and AST validations. **a** Heatmap with ARG count. The barplots illustrate the number of ARGs across rows and across columns. Species were identified using BLAST. Only ARGs with probability > 0.8 were considered, as recommended. **b** Antibacterial susceptibility tests (AST) on *E. bugandensis* and *B. cereus* strains for several antibiotics (top), and comparison with machine learning predictions shown in **(a)** (bottom). **c** Scatterplot of zone of inhibition value (in mm.) and ARG count shown in **(b)**, together with a linear model fit. Pearson's product-moment correlation values are indicated

To experimentally validate machine learning predictions on previously unobserved AMR patterns above, we performed Antibacterial Susceptibility Tests (AST) for the species found to be potentially most pathogenic, in our case *E. bugandensis* and *B. cereus* as they have a higher number of ARGs (Table [S1](#); Fig. [4a](#)). For that, we use disc diffusion on strains isolated at the ISS for the following antimicrobials: Cefazolin (beta-lactam), Cefoxitin (beta-lactam), Ciprofloxacin (quinolone), Erythromycin (MLS), Gentamycin (aminoglycoside), Oxacillin (beta-lactam), Penicillin (beta-lactam), Rifampin (rifamycin), and Tobramycin (aminoglycoside) (Fig. [4b](#)). The prediction patterns closely matched the AST results (Fig. [4b](#)), although DeepARG failed to detect Rifampin resistance, especially for *E. bugandensis*.

Although different antibiotics have different inhibitory zone cut-offs for a strain to be considered as resistant (Table [S2](#)), remarkably we found an inverse correlation between the zone of inhibition and ARG count for *B. cereus* ($r = -0.637$, Pearson's product-moment correlation, $p = 2.2e-7$) and *E. bugandensis* ($r = -0.517$; $p = 0.0002765$) (Fig. [4c](#)), demonstrating the applicability and high accuracy of computational prediction of AMR for microbiome data obtained in space.

Discussion

Many ARGs that present high probability but low sequence identity to known sequences will be missed using traditional 'best-hit' approaches that require a high degree of sequence identity. To solve this, computational methods have been developed to identify AMR in genomes and metagenomes [[3](#), [9](#), [16](#), [33](#), [41](#)]. Despite these developments, a consensus approach to detect AMR in metagenomics datasets is yet to be defined [[41](#)]. The objective of this study was to identify the AMR genes associated with cultivated strains and metagenomes generated from the ISS environmental surfaces using an accurate deep learning approach (Fig. [1](#)).

Firstly, we re-analyzed shotgun metagenome sequences of 21 environmental samples that were treated with PMA (viable microbes), and their associated 24 MAGs retrieved from the PMA-treated samples. The re-analysis showed increased read counts associated with AMR and in more locations for flight 3 when considering MAGs (Fig. [2](#)). This

could be explained due to the ISS crew being replaced during Flight 3. The abundance of *Enterobacteriaceae* in Flight 3 was discussed in Singh et al. [43]. We have not observed any differences between early vs. late ISS microbiome cultures. For example, *Enterobacter bugandensis* strains were isolated from F2 and F3 sample sets, but the genome comparison and phenotype analyses revealed limited change, with a maximum of 15 SNPs among ISS isolates [44]. In addition, *K. piersonii* spread across four different locations (L1, L5, L7, L8) at Flight 3, presenting resistance to specific antibiotics (glyco/peptide, fluoroquinolone and MLS) (Fig. 3c). We have previously isolated strains from Locations 1, 2, 5, 6, and 7, defining a novel bacterial genus from the ISS samples [45]. While *K. piersonii* do have virulence genes in the genome, a dichotomy was found as disc diffusion tests revealed multi-drug resistance, while the PathogenFinder algorithm predicted *K. piersonii* strains as non-human pathogens. All seven *K. piersonii* isolates were resistant to cefoxitin (beta_lactam class in DeepARG), erythromycin (MLS), oxacillin (beta_lactam), penicillin (beta_lactam), and rifampin. At the same time, all strains were susceptible to cefazolin, ciprofloxacin (quinolone), and tobramycin (aminoglycoside) [45]. The DeepARG database does not include some of these antibiotics, but we found AMR sequences related to resistance to (glyco)peptide, fluoroquinolone, and MLS, validating some previous results. Therefore, PathogenFinder [17] results in Singh et al. [45] suggesting *K. piersonii* as a non-human pathogen should be treated with caution. Furthermore, the strain YU22 (closest match is IIIF1SW-P2^T detected as ISS) isolated in urine microbiome of a kidney stone patient has shown to be an uropathogenic bacteria, showing many virulence factors that are needed for host cell invasion and colonization [40].

Secondly, the whole-genome sequences (WGS) of 226 pure strains (cultivable microbes) were analyzed to identify AMR genes (Fig. 4a). We found the human pathogens *E. bugandensis* and *B. cereus* presenting many potential ARGs in the MT-1 scaffolds. Up to five strains isolated from the ISS have been closely related to the type strain EB-247^T and two clinical isolates (153_ECLO and MBRL 1077) and share similar AMR patterns [44]. One hundred twelve genes were found to be involved in virulence, disease, and defence in the ISS strains [44]. Our re-analysis confirms the *multi-drug* resistance (MDR) to antibiotics for the ISS isolates, which is the highest among all the isolates. Our previous research uncovered the presence of genes associated with MDR efflux pump [44], belonging to RND (resistance, nodulation and cell division) protein family, which are reported to be the major contributors of resistance to antibiotic and other toxic compounds to the bacteria [20]. MDR has been reported to play role in the physiological function and confer resistance to substances like host defense molecule and bile, which can lead to pathogenicity in humans [48]. Unlike in Singh et al. [43], we found fluoroquinolone resistance low, and null for fosfomycin. Conversely, *B. cereus* is a gram-positive bacterium commonly found in food. After infection, most emetic patients recover within 24 hours, but in some cases, the toxin can be fatal via a fulminant hepatic failure [22, 34]. Overall, *multi-drug* resistance was found widespread in many microbes.

Third, phenotypic antibiotic resistance testing data obtained from traditional antibiotic tests generated for biosafety level 2 strains were compared with the computational approaches that predicted the presence of the AMR genes, showing an excellent agreement for the antibiotics tested (Fig. 4b, c). A disadvantage of the deep learning model developed by Arango-Argoty et al. [3] is that the prediction can disentangle the family of antibiotics but not specific compounds.

Many studies have shown the association between several microorganisms (bacterial, as well as phage and non-phase viral sequences) and several cancer features. Although it is unclear whether this corresponds to correlation or causation, the microbiome can undoubtedly be used as a cancer biomarker. For instance, certain strains of *Fusobacterium* sp. can be utilized as an independent diagnostic assay for colon cancer [62]. Therefore, a better understanding of the microbial communities and their degree of pathogenicity in surface-human microbiomes in space could also be useful for human health monitorization with detection and prognostic values in long term space travel. Rather than the gather-and-return sampling model currently used for the ISS, new developments in sequencing technologies in combination with Artificial Intelligence will allow for efficient analysis onboard the ISS and in long-duration space missions.

We are currently collecting more data for Microbial Tracking-2 (MT-2) and MT-3 missions. We plan to extend the AMR catalog, characterize microbial diversity, and monitor the evolution of AMR in longer time periods to discover new factors involved in pathogenicity of microorganisms exposed to space conditions.

Methods

Metagenome-Assembled Genomes (MAGs) methodology

The paired-end 100-bp metagenomic reads from NCBI Short Read Archive (SRA) under the bio-project number PRJNA438545 were processed with Trimmomatic [12] to trim adapter sequences and low-quality ends, with a minimum Phred score of 20 across the entire length of the read used as a quality cut-off. Reads shorter than 80 bp were removed after trimming. Remaining high-quality reads were subsequently assembled using metaSPAdes [37]. Contigs were binned using Metabat2 v2.11.3 [31]. Recovered genomes were evaluated with CheckM [39], and a recovered genome was considered good with at least 90% completeness and at most 10% contamination. Each genome was subsequently annotated with the help of Rapid Annotations using Subsystems Technology (RAST), and near identifications were predicted [6].

Sample collection for Microbial Tracking 1 mission

During the microbial tracking investigation to characterize airborne and surface-associated microbial population aboard the International Space Station, samples were collected from ISS locations, and ground samples were collected from the Crew resupply vehicle. A sterile polyester wipe premoistened with Phosphate buffered saline (PBS) was used to collect the samples from various areas across these ISS locations, and the details and description of the sample collection and cultivation have already been reported in Checinska Sielaff et al. [14] and Singh et al. [43]. Pure isolates were selected and sub-cultured, and the sub-cultures were sequenced.

Isolates from Microbial Tracking 1 mission

To create the whole-genome sequences (WGS) of these strains, shotgun libraries were prepared using the Illumina Nextera Flex protocol [44], using NovaSeq 6000 S4 flow cell 2150 paired-end (PE) sequencing. Verification of the quality of the raw sequencing data was carried out using FastQC v0.11.7 [2]. Quality control for adapter trimming and quality filtering were performed using fastp v0.20.0 [15], and then SPAdes v3.11.1 [7] was used to assemble all the cleaned sequences. Fastp quality control was based on the following three parameters: (i) correction of mismatches in overlapped regions of paired-end reads, (ii) trimming of autodetected adapter sequences, and (iii) quality trimming at the 5' and 3' ends. To determine the quality of the assembled sequences, the number of contigs, the N50 value, and the total length were calculated using QUAST v5.0.2 [26]. Default parameters were used for all software. The average nucleotide identity (ANI) [59] was calculated using OrthoANIu by comparing each of the scaffolds to the WGS of the respective type strains.

Identification of ORFs in microbial DNA sequences

Glimmer (Gene Locator and Interpolated Markov ModelER) v3.02 was used with default parameters to identify the coding regions and distinguish them from non-coding DNA in MAGs and MT-1 scaffolds that could be used as an input in DeepARG-LS. Minimum gene length was indicated as 50 bp ('glimmer3 -g50'). Glimmer reads DNA sequences in a FASTA file format and predicts genes in them using an Interpolated Context Model [21].

Prediction of antibiotic resistance genes in short reads and full-gene length sequences

DeepARG version 2 [3], a deep learning-based approach for predicting ARGs and annotation, was run with the '--reads' option (DeepARG-SS) for NGS reads and the '--genes' option (DeepARG-LS) for longer gene-like sequences obtained with Glimmer. The DeepARG model consists of four dense hidden layers of 2000, 1000, 500, and 100 units that propagate a bit score distribution. The output layer of the deep neural network consists of 30 units that correspond to the antibiotic resistance categories (102 antibiotics consolidated into 30 antibiotic categories). The model was trained with a curated database of 14,933 genes from three databases (CARD, ARDB, and UNIPROT) [3]. Default options were used: 50% minimum percentage of identity to consider, significance of the prediction probability cut-off of 0.8 as recommended [3], and E-value of alignments (default $1e-10$). The software was downloaded from <https://bitbucket.org/gusphdproj/deeparg-ss>.

Microbial nucleotide BLAST

Nucleotide-Nucleotide BLAST 2.10.1+ (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>) was used to identify microbial species associated to MT-1 scaffolds. Sequences producing significant alignments were ranked and the species associated to maximum Score (bits) and minimum *E* value was deemed as the closest match.

Gene annotations and pan-genome analysis of MT-1 scaffolds

A Docker image with the software tool Prokka v1.14.5 [42] was pulled and used with default parameters to annotate bacterial isolates. The master annotations in GFF3 format for the 226 isolates have been deposited at the Zenodo platform [DOI: 10.5281/zenodo.6518836]. Using the annotated assemblies in GFF3 format for *E. bugandensis* and *B. cereus* strains, Roary v3.13.0 was run with default parameters to compare isolates based on genes and the number of isolates they are present in.

Phenotypic antibiotic resistance testing

Disc assays experiments were performed and reported in Urbaniak et al. [51]. The isolates were streaked from glycerol stocks onto R2A plates. A single colony was inoculated into 5 mL Tryptic Soy Broth (TSB) and grown overnight at 30°C. Aliquots of 100 µL were plated on TSA. Agar diffusion discs (BD BBLTM Sensi-Disc™, Franklin Lakes, NJ) were placed aseptically on a plate and the strains were incubated at 37 °C for 24 h. The tested antibiotics included: 30-µg cefazolin (CZ-30); 30-µg cefoxitin (FOX-30), 5 µg ciprofloxacin (CIP-5), 15 µg erythromycin (E-15), 10-µg gentamicin (GM-10), 1-µg oxacillin (OX), 10-µg penicillin (P-10), 5-µg rifampin (RA-5), and 10 µg tobramycin (NN-10). The diameter of inhibition zones was measured for each antibiotic disk and recorded in millimeters. The resistance results were compared with the zone diameter interpretive charts provided by the manufacturer. When the spontaneous mutants were present in response to some antibiotics, they were isolated, subcultured and tested for the specific antibiotic resistance.

Data availability

Raw metagenomics reads from three flights on multiple locations were downloaded from NASA GeneLab repository <https://genelab.nasa.gov> (GLDS-69). The ISS MAGs datasets are available at the SRA database under the accession number: PRJNA438545 [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA438545/>]. Microbial tracking-1 (MT-1) datasets (raw sequencing data and sequence assembly files) were obtained from GLDS-67, GLDS-302, GLDS-303, GLDS-309, GLDS-311 and GLDS-350. The rest of the samples are deposited at DDBJ/ ENA/GenBank or are unpublished.

Supplementary Information

[40168_2022_1332_MOESM1_ESM.pdf](#) (50.1KB, pdf)

Additional file 1: Figure S1. Distribution of DeepARG classification probability and best-hit identity in MT-1 pure strains isolated from the ISS. The blue dashed line indicates 50% sequence identity.

[40168_2022_1332_MOESM2_ESM.pdf](#) (100.3KB, pdf)

Additional file 2: Figure S2. Gene annotations and pan-genome analysis of MT-1 scaffolds. (a) Scatterplots of number of coding sequences (CDS) found by Prokka, genome sizes, and Antibiotic resistance genes (ARGs) detected by DeepARG. (b) Frequency of genes in the core and accessory groups (soft core, shell, cloud) for ISS isolates of *E. bugandensis* (10 strains) and *B. cereus* (10 strains).

[40168_2022_1332_MOESM3_ESM.xlsx](#) (16.6KB, xlsx)

Additional file 3: Table S1. Rank of MT-1 isolates, ordered by number of ARGs predicted, shown in Fig. 4a. Species information obtained from Microbial Nucleotide BLAST.

[40168_2022_1332_MOESM4_ESM.xlsx](#) (15.6KB, xlsx)

Additional file 4: Table S2. Phenotypic antibiotic resistance testing results for *E. bugandensis* and *B. cereus*.

[40168_2022_1332_MOESM5_ESM.zip](#) (73.1MB, zip)

Additional file 5: File S1. Raw results of the DeepARG analyses (zip compressed).

Acknowledgements

We thank Dr. Sylvain Costes, Dr. Jonathan Galazka, and Dr. Daniel C. Berrios for the initial conversations that inspired this project. The authors would like to acknowledge the members of the Microbiome and Multi-Omics/Systems Biology Analysis Working Groups of NASA GeneLab. Part of the research described in this manuscript was performed at the Jet Propulsion Laboratory, California Institute of Technology under a contract with NASA. We would like to thank Microbial Tracking -1 and 2 members for isolating the strain and generating draft assembly of genomes. We thank Biotechnology and Planetary Protection Group members for supporting sample analyses. We also acknowledge the Jet Propulsion Laboratory supercomputing facility staff, notably Narendra J. Patel (Jimmy) and Edward Villanueva, for their continuous support in providing the best possible infrastructure for BIG-DATA analysis.

Abbreviations

AMR

Antimicrobial resistance

ARG

Antibiotic resistance gene

AST

Antibacterial susceptibility test

ISS

International Space Station

MAG

Metagenome-assembled genome

MDR

Multi-drug resistance

MT-1

Microbial Tracking–1

PMA

Propidium monoazide

WGS

Whole-genome sequencing

Authors' contributions

PM performed the computational analyses and led the writing. NKS, KV, JMW, and CEM provided data, exchange of ideas, and edits to the manuscript. EG and FHdO contributed to machine learning aspects. AB supervised and supported

the study, wrote and provided final approval of the manuscript. All authors provided feedback and contributed to the research and final manuscript. All authors read and approved the final manuscript.

Funding

AB was supported and funded by NASA grant 16-ROSBFP_GL-0005: NNH16ZTT001N-FG Appendix G: Solicitation of Proposals for Flight and Ground Space Biology Research (Award Number: 80NSSC19K0883) and The Translational Research Institute for Space Health through NASA Cooperative Agreement NNX16AO69A (T-0404). This work was partially supported by the ESA Space Omics Topical Team, funded by the ESA grant/contract 4000131202/20/NL/PG/pt “Space Omics: Towards an integrated ESA/NASA –omics database for spaceflight and ground facilities experiments”. CEM was supported by NASA grants (NNX16AO69A, NNX14AH50G, NNX17AB26G), the NIH (R01AI151059, U01DA053941) and Igor Tulchinsky and the WorldQuant Foundation. This research was funded by a 2012 Space Biology NNH12ZTT001N grant no. 19-12829-26 under Task Order NNN13D111T award to KV, which also funded the post-doctoral fellowships for NKS.

Availability of data and materials

The code use for the analysis is available at <https://github.com/pmb59/AMRISS> .

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors have read and agreed with the manuscript.

Competing interests

The authors declare that they have no competing interests.

Footnotes

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Kasthuri Venkateswaran and Afshin Beheshti are joint senior authors.

References

1. Afshinnekoo E, Scott RT, MacKay MJ, Pariset E, Cekanaviciute E, Barker R, Gilroy S, Hassane D, Smith SM, Zwart SR, Nelman-Gonzalez M, Crucian BE, Ponomarev SA, Orlov OI, Shiba D, Muratani M, Yamamoto M, Richards SE, Vaishampayan PA, Meydan C, Foox J, Myrrhe J, Istasse E, Singh N, Venkateswaran K, Keune JA, Ray HE, Basner M, Miller J, Vitaterna MH, Taylor DM, Wallace D, Rubins K, Bailey SM, Grabham P, Costes SV, Mason CE, Beheshti A. Fundamental Biological Features of Spaceflight: Advancing the Field to Enable Deep-Space Exploration. *Cell*. 2020;183(5):1162–1184. doi: 10.1016/j.cell.2020.10.050. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
2. Andrews S. FastQC: a quality tool for high throughput sequence data. 2015. [[Google Scholar](#)]
3. Arango-Argoty G, Garner E, Pruden A, Heath LS, Vikesland P, Zhang L. DeepARG: a deep learning approach for predicting antibiotic resistance genes from metagenomic data. *Microbiome*. 2018;6(1):23. doi: 10.1186/s40168-018-0401-z. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
4. Aunins TR, Erickson KE, Prasad N, et al. Spaceflight modifies escherichia coli gene expression in response to antibiotic exposure and reveals role of oxidative stress response. *Front Microbiol*. 2018;9:310. doi: 10.3389/fmicb.2018.00310. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
5. Avila-Herrera A, Thissen J, Urbaniak C, et al. Crewmember microbiome may influence microbial composition of ISS habitable surfaces. *PLoS One*. 2020;15(4):e0231838. doi: 10.1371/journal.pone.0231838. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
6. Aziz RK, Bartels D, Best AA, Dejongh M, Disz T, Edwards RA, Formsma K, Gerdes S, Glass EM, Kubal M, Meyer F, Olsen GJ, Olson R, Osterman AL, Overbeek RA, Mcneil LK, Paarmann D, Paczian T, Parrello B, Pusch GD, Reich C, Stevens R, Vassieva O, Vonstein V, Wilke A, Zagnitko O. The RAST Server: rapid annotations using subsystems technology. *BMC Genomics*. 2008;9:75. doi: 10.1186/1471-2164-9-75. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]
7. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Pribelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA. SPAdes: a new

genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol.* 2012;19:455–477. doi: 10.1089/cmb.2012.0021. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

8. Be NA, Avila-Herrera A, Allen JE, et al. Whole metagenome profiles of particulates collected from the International Space Station. *Microbiome.* 2017;5(1):81. doi: 10.1186/s40168-017-0292-4. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

9. Berglund F, Österlund T, Boulund F, Marathe NP, Larsson DGJ, Kristiansson E. Identification and reconstruction of novel antibiotic resistance genes from metagenomes. *Microbiome.* 2019;7(1):52. doi: 10.1186/s40168-019-0670-1. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

10. Bijlani S, Singh NK, Mason CE, Wang CCC, Venkateswaran K. Draft Genome Sequences of *Sphingomonas* Species Associated with the International Space Station. *Microbiol Resour Announc.* 2020;9(25):e00578–e00520. doi: 10.1128/MRA.00578-20. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

11. Bijlani S, Singh NK, Mason CE, Wang CCC, Venkateswaran K. Draft Genome Sequences of *Tremellomycetes* Strains Isolated from the International Space Station. *Microbiol Resour Announc.* 2020;9(26):e00504–e00520. doi: 10.1128/MRA.00504-20. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

12. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30:2114–2120. doi: 10.1093/bioinformatics/btu170. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

13. Boolchandani M, D'Souza AW, Dantas G. Sequencing-based methods and resources to study antimicrobial resistance. *Nat Rev Genet.* 2019;20(6):356–370. doi: 10.1038/s41576-019-0108-4. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

14. Checinska Sielaff A, Urbaniak C, Mohan GBM, Stepanov VG, Tran Q, Wood JM, Minich J, McDonald D, Mayer T, Knight R, Karouia F, Fox GE, Venkateswaran K. Characterization of the total and viable bacterial and fungal communities associated with the International Space Station surfaces. *Microbiome.* 2019;7(1):50. doi: 10.1186/s40168-019-0666-x. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

15. Chen S, Zhou Y, Chen Y, Gu J. fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics.* 2018;34:i884–i890. doi: 10.1093/bioinformatics/bty560. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

16. Chowdhury AS, Call DR, Broschat SL. Antimicrobial resistance prediction for gram-negative bacteria via game theory-based feature evaluation [published correction appears in *Sci Rep.* 2020 Jan 30;10(1):1846] *Sci Rep.* 2019;9(1):14487. doi: 10.1038/s41598-019-50686-z. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google](#)

17. Cosentino S, Voldby Larsen M, Møller Aarestrup F, Lund O. PathogenFinder--distinguishing friend from foe using bacterial whole genome sequence data [published correction appears in PLoS One. 2013;8(12). doi:10.1371/annotation/b84e1af7-c127-45c3-be22-76abd977600f] PLoS One. 2013;8(10):e77302. doi: 10.1371/journal.pone.0077302. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

18. Danko D, Bezdan D, Afshin EE, Ahsanuddin S, Bhattacharya C, Butler DJ, Chng KR, Donnellan D, Hecht J, Jackson K, Kuchin K, Karasikov M, Lyons A, Mak L, Meleshko D, Mustafa H, Mutai B, Neches RY, Ng A, Nikolayeva O, Nikolayeva T, Png E, Ryon KA, Sanchez JL, Shaaban H, Sierra MA, Thomas D, Young B, Abudayyeh OO, Alicea J, Bhattacharyya M, Blekhman R, Castro-Nallar E, Cañas AM, Chatzieftimiou AD, Crawford RW, De Filippis F, Deng Y, Desnues C, Dias-Neto E, Dybwad M, Elhaik E, Ercolini D, Frolova A, Gankin D, Gootenberg JS, Graf AB, Green DC, Hajirasouliha I, Hastings JJA, Hernandez M, Iraola G, Jang S, Kahles A, Kelly FJ, Knights K, Kyrpides NC, Łabaj PP, Lee PKH, Leung MHY, Ljungdahl PO, Mason-Buck G, McGrath K, Meydan C, Mongodin EF, Moraes MO, Nagarajan N, Nieto-Caballero M, Noushmehr H, Oliveira M, Ossowski S, Osuolale OO, Özcan O, Paez-Espino D, Rascovan N, Richard H, Rätsch G, Schriml LM, Semmler T, Sezerman OU, Shi L, Shi T, Siam R, Song LH, Suzuki H, Court DS, Tighe SW, Tong X, Udekwu KI, Ugalde JA, Valentine B, Vassilev DI, Vayndorf EM, Velavan TP, Wu J, Zambrano MM, Zhu J, Zhu S, Mason CE. International MetaSUB Consortium. A global metagenomic map of urban microbiomes and antimicrobial resistance. Cell. 2021;184(13):3376–3393. doi: 10.1016/j.cell.2021.05.002. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

19. Danko DC, Singh N, Butler DJ, Mozsary C, Jiang P, Keshavarzian A, et al. Genetic and immunological evidence for microbial transfer between the international space station and an astronaut. bioRxiv. 2020. 11.10.376954. 10.1101/2020.11.10.376954.

20. Daury L, Orange F, Taveau JC, Verchere A, Monlezun L, Gounou C, Marreddy RK, Picard M, Broutin I, Pos KM, et al. Tripartite assembly of RND multidrug efflux pumps. Nat Commun. 2016;7:10731. doi: 10.1038/ncomms10731. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

21. Delcher AL, Bratke KA, Powers EC, Salzberg SL. Identifying bacterial genes and endosymbiont DNA with Glimmer. Bioinformatics. 2007;23(6):673–679. doi: 10.1093/bioinformatics/btm009. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

22. Dierick K, Van Coillie E, Swiecicka I, Meyfroidt G, et al. Fatal family outbreak of *Bacillus cereus*-associated food poisoning. J Clin Microbiol. 2005;43(8):4277–4279. doi: 10.1128/JCM.43.8.4277-4279.2005. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

23. Du B, Daniels VR, Vaksman Z, Boyd JL, Crady C, Putcha L. Evaluation of physical and chemical changes in pharmaceuticals flown on space missions. AAPS J. 2011;13(2):299–308. doi: 10.1208/s12248-011-9270-0.

[\[DOI\]](#) [\[PMC free article\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

24. Fukuda T, Fukuda K, Takahashi A, et al. Analysis of deletion mutations of the rpsL gene in the yeast *Saccharomyces cerevisiae* detected after long-term flight on the Russian space station Mir. *Mutat Res*. 2000;470(2):125–132. doi: 10.1016/s1383-5742(00)00054-5. [\[DOI\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

25. Garrett-Bakelman FE, Darshi M, Green SJ, Gur RC, Lin L, Macias BR, McKenna MJ, Meydan C, Mishra T, Nasrini J, Piening BD, Rizzardi LF, Sharma K, Siamwala JH, Taylor L, Vitaterna MH, Afkarian M, Afshinnekoo E, Ahadi S, Ambati A, Arya M, Bezdán D, Callahan CM, Chen S, Choi AMK, Chlipala GE, Contrepois K, Covington M, Crucian BE, De Vivo I, Dinges DF, Ebert DJ, Feinberg JI, Gandara JA, George KA, Goutsias J, Grills GS, Hargens AR, Heer M, Hillary RP, Hoofnagle AN, Hook VYH, Jenkinson G, Jiang P, Keshavarzian A, Laurie SS, Lee-McMullen B, Lumpkins SB, MacKay M, Maienschein-Cline MG, Melnick AM, Moore TM, Nakahira K, Patel HH, Pietrzyk R, Rao V, Saito R, Salins DN, Schilling JM, Sears DD, Sheridan CK, Stenger MB, Tryggvadottir R, Urban AE, Vaisar T, Van Espen B, Zhang J, Ziegler MG, Zwart SR, Charles JB, Kundrot CE, Scott GBI, Bailey SM, Basner M, Feinberg AP, Lee SMC, Mason CE, Mignot E, Rana BK, Smith SM, Snyder MP, Turek FW. The NASA twins study: a multidimensional analysis of a year-long human spaceflight. *Science*. 2019;364(6436):eaau8650. doi: 10.1126/science.aau8650. [\[DOI\]](#) [\[PMC free article\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

26. Gurevich A, Saveliev V, Vyahhi N, Tesler G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics*. 2013;29:1072–1075. doi: 10.1093/bioinformatics/btt086. [\[DOI\]](#) [\[PMC free article\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

27. Hadjadj L, Baron SA, Diene SM, Rolain JM. How to discover new antibiotic resistance genes? *Expert Rev Mol Diagn*. 2019;19(4):349–362. doi: 10.1080/14737159.2019.1592678. [\[DOI\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

28. Hammond TG, Stodieck L, Birdsall HH, et al. Effects of microgravity on the virulence of *Listeria monocytogenes*, *Enterococcus faecalis*, *Candida albicans*, and methicillin-resistant *Staphylococcus aureus*. *Astrobiology*. 2013;13(11):1081–1090. doi: 10.1089/ast.2013.0986. [\[DOI\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

29. Hendriksen RS, Bortolaia V, Tate H, Tyson GH, Aarestrup FM, McDermott PF. Using Genomics to Track Global Antimicrobial Resistance. *Front Public Health*. 2019;7:242. doi: 10.3389/fpubh.2019.00242. [\[DOI\]](#) [\[PMC free article\]](#) [\[PubMed\]](#) [\[Google Scholar\]](#)

30. Juergensmeyer MA, Juergensmeyer EA, Guikema JA. Long-term exposure to spaceflight conditions affects bacterial response to antibiotics. *Microgravity Sci Technol*. 1999;12(1):41. [\[PubMed\]](#) [\[Google Scholar\]](#)

31. Kang DD, Froula J, Egan R, Wang Z. MetaBAT, an efficient tool for accurately reconstructing single

genomes from complex microbial communities. PeerJ. 2015;3:e1165. doi: 10.7717/peerj.1165. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

32. Kim W, Tengra FK, Young Z, et al. Spaceflight promotes biofilm formation by *Pseudomonas aeruginosa*. PLoS One. 2013;8(4):e62437. doi: 10.1371/journal.pone.0062437. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

33. Lakin SM, Kuhnle A, Alipanahi B, et al. Hierarchical Hidden Markov models enable accurate and diverse detection of antimicrobial resistance sequences. Commun Biol. 2019;2:294. doi: 10.1038/s42003-019-0545-9. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

34. Mahler H, Pasi A, Kramer JM, Schulte P, et al. Fulminant liver failure in association with the emetic toxin of *Bacillus cereus*. N Engl J Med. 1997;336(16):1142–1148. doi: 10.1056/NEJM199704173361604. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

35. Morrison MD, Thissen JB, Karouia F, Mehta S, Urbaniak C, Venkateswaran K, Smith DJ, Jaing C. Investigation of Spaceflight Induced Changes to Astronaut Microbiomes. Front Microbiol. 2021;12:659179. doi: 10.3389/fmicb.2021.659179. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

36. Nickerson CA, Ott CM, Wilson JW, Ramamurthy R, Pierson DL. Microbial responses to microgravity and other low-shear environments. Microbiol Mol Biol Rev. 2004;68(2):345–361. doi: 10.1128/MMBR.68.2.345-361.2004. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

37. Nurk S, Meleshko D, Korobeynikov A, Pevzner PA. metaSPAdes: a new versatile metagenomic assembler. Genome Res. 2017;27:824–834. doi: 10.1101/gr.213959.116. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

38. Page AJ, Cummins CA, Hunt M, Wong VK, Reuter S, Holden MT, Fookes M, Falush D, Keane JA, Parkhill J. Roary: rapid large-scale prokaryote pan genome analysis. Bioinformatics. 2015;31(22):3691–3693. doi: 10.1093/bioinformatics/btv421. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

39. Parks DH, Imelfort M, Skennerton CT, Hugenholtz P, Tyson GW. CheckM: assessing the quality of microbial genomes recovered from isolates, single cells, and metagenomes. Genome Res. 2015;25:1043–1055. doi: 10.1101/gr.186072.114. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

40. Rekha PD, Hameed A, Manzoor MAP, Suryavanshi MV, Ghate SD, Arun AB, Rao SS, Bajire SK, Mujeeburaheeman M, Young CC. First Report of Pathogenic Bacterium *Kalamiella piersonii* Isolated from Urine of a Kidney Stone Patient: Draft Genome and Evidence for Role in Struvite Crystallization. Pathogens. 2020;9(9):711. doi: 10.3390/pathogens9090711. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

41. Ruppé E, Ghoulane A, Tap J, et al. Prediction of the intestinal resistome by a three-dimensional structure-

based method. *Nat Microbiol.* 2019;4(1):112–123. doi: 10.1038/s41564-018-0292-6. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

42. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics.* 2014;30(14):2068–2069. doi: 10.1093/bioinformatics/btu153. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

43. Singh NK, Wood JM, Karouia F, Venkateswaran K. Succession and persistence of microbial communities and antimicrobial resistance genes associated with International Space Station environmental surfaces. *Microbiome.* 2018;6(1):204. doi: 10.1186/s40168-018-0585-2. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

44. Singh NK, Bezdan D, Checinska Sielaff A, Wheeler K, Mason CE, Venkateswaran K. Multi-drug resistant *Enterobacter bugandensis* species isolated from the International Space Station and comparative genomic analyses with human pathogenic strains. *BMC Microbiol.* 2018;18:175. doi: 10.1186/s12866-018-1325-2. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

45. Singh NK, Wood JM, Mhatre SS, Venkateswaran K. Metagenome to phenome approach enables isolation and genomics characterization of *Kalamiella piersonii* gen. nov., sp. nov. from the International Space Station. *Appl Microbiol Biotechnol.* 2019;103(11):4483–4497. doi: 10.1007/s00253-019-09813-z. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

46. Sonnenfeld G, Shearer WT. Immune function during space flight. *Nutrition.* 2002;18(10):899–903. doi: 10.1016/s0899-9007(02)00903-6. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

47. Su M, Satola SW, Read TD. Genome-based prediction of bacterial antibiotic resistance. *J Clin Microbiol.* 2019;57(3):e01405–e01418. doi: 10.1128/JCM.01405-18. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

48. Sun J, Deng Z, Yan A. Bacterial multidrug efflux pumps: mechanisms, physiology and pharmacological exploitations. *Biochem Biophys Res Commun.* 2014;453(2):254–267. doi: 10.1016/j.bbrc.2014.05.090. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

49. Taylor PW. Impact of space flight on bacterial virulence and antibiotic susceptibility. *Infect Drug Resist.* 2015;8:249–262. doi: 10.2147/IDR.S67275. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

50. Tixador R, Richoille G, Gasset G, et al. Study of minimal inhibitory concentration of antibiotics on bacteria cultivated in vitro in space (Cytos 2 experiment) *Aviat Space Environ Med.* 1985;56(8):748–751. [[PubMed](#)] [[Google Scholar](#)]

51. Urbaniak C, Sielaff AC, Frey KG, et al. Detection of antimicrobial resistance genes associated with the International Space Station environmental surfaces. *Sci Rep.* 2018;8(1):814. doi: 10.1038/

s41598-017-18506-4. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

52. Urbaniak C, Lorenzi H, Thissen J, et al. The influence of spaceflight on the astronaut salivary microbiome and the search for a microbiome biomarker for viral reactivation. *Microbiome*. 2020;8(1):56. doi: 10.1186/s40168-020-00830-z. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

53. Venkateswaran K, Vaishampayan P, Cisneros J, Pierson DL, Rogers SO, Perry J. International Space Station environmental microbiome - microbial inventories of ISS filter debris. *Appl Microbiol Biotechnol*. 2014;98(14):6453–6466. doi: 10.1007/s00253-014-5650-6. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

54. Venkateswaran K, Singh NK, Checinska Sielaff A, et al. Non-Toxin-Producing *Bacillus cereus* Strains Belonging to the *B. anthracis* Clade Isolated from the International Space Station. *mSystems*. 2017;2(3):e00021–e00017. doi: 10.1128/mSystems.00021-17. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

55. Voorhies AA, Mark Ott C, Mehta S, et al. Study of the impact of long-duration space missions at the International Space Station on the astronaut microbiome. *Sci Rep*. 2019;9(1):9911. doi: 10.1038/s41598-019-46303-8. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

56. Wilson JW, Ott CM, Höner zu Bentrup K, et al. Space flight alters bacterial gene expression and virulence and reveals a role for global regulator Hfq. *Proc Natl Acad Sci U S A*. 2007;104(41):16299–16304. doi: 10.1073/pnas.0707155104. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

57. Wissel EF, Talbot BM, Johnson BA, Petit RA, Hertzberg V, Dunlop A, et al. Benchmarking software to predict antibiotic resistance phenotypes in shotgun metagenomes using simulated data. *bioRxiv*. 2022.01.13.476279. 10.1101/2022.01.13.476279.

58. World Health Organization . Global Action Plan on Antimicrobial Resistance. 2015. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

59. Yoon S-H, Ha S-M, Lim J, Kwon S, Chun J. A large-scale evaluation of algorithms to calculate average nucleotide identity. *Antonie Van Leeuwenhoek*. 2017;110:1281–1286. doi: 10.1007/s10482-017-0844-4. [[DOI](#)] [[PubMed](#)] [[Google Scholar](#)]

60. Zankari E, Hasman H, Cosentino S, et al. Identification of acquired antimicrobial resistance genes. *J Antimicrob Chemother*. 2012;67:2640–2644. doi: 10.1093/jac/dks261. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

61. Zea L, Larsen M, Estante F, et al. Phenotypic Changes Exhibited by *E. coli* Cultured in Space. *Front Microbiol*. 2017;8:1598. doi: 10.3389/fmicb.2017.01598. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

62. Zhang X, et al. Fecal fusobacterium nucleatum for the diagnosis of colorectal tumor: a systematic review and meta-analysis. *Cancer Med.* 2019;8:480–491. doi: 10.1002/cam4.1850. [[DOI](#)] [[PMC free article](#)] [[PubMed](#)] [[Google Scholar](#)]

Associated Data

This section collects any data citations, data availability statements, or supplementary materials included in this article.

Supplementary Materials

[40168_2022_1332_MOESM1_ESM.pdf](#) (50.1KB, pdf)

Additional file 1: Figure S1. Distribution of DeepARG classification probability and best-hit identity in MT-1 pure strains isolated from the ISS. The blue dashed line indicates 50% sequence identity.

[40168_2022_1332_MOESM2_ESM.pdf](#) (100.3KB, pdf)

Additional file 2: Figure S2. Gene annotations and pan-genome analysis of MT-1 scaffolds. (a) Scatterplots of number of coding sequences (CDS) found by Prokka, genome sizes, and Antibiotic resistance genes (ARGs) detected by DeepARG. (b) Frequency of genes in the core and accessory groups (soft core, shell, cloud) for ISS isolates of *E. bugandensis* (10 strains) and *B. cereus* (10 strains).

[40168_2022_1332_MOESM3_ESM.xlsx](#) (16.6KB, xlsx)

Additional file 3: Table S1. Rank of MT-1 isolates, ordered by number of ARGs predicted, shown in Fig. [4a](#). Species information obtained from Microbial Nucleotide BLAST.

[40168_2022_1332_MOESM4_ESM.xlsx](#) (15.6KB, xlsx)

Additional file 4: Table S2. Phenotypic antibiotic resistance testing results for *E. bugandensis* and *B. cereus*.

[40168_2022_1332_MOESM5_ESM.zip](#) (73.1MB, zip)

Additional file 5: File S1. Raw results of the DeepARG analyses (zip compressed).

Data Availability Statement

Raw metagenomics reads from three flights on multiple locations were downloaded from NASA GeneLab repository <https://genelab.nasa.gov> (GLDS-69). The ISS MAGs datasets are available at the SRA database under the accession number: PRJNA438545 [<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA438545/>]. Microbial tracking-1 (MT-1) datasets (raw sequencing data and sequence assembly files) were obtained from GLDS-67, GLDS-302, GLDS-303, GLDS-309, GLDS-311 and GLDS-350. The rest of the samples are deposited at DDBJ/ ENA/GenBank or are unpublished.

The code use for the analysis is available at <https://github.com/pmb59/AMRISS> .

Articles from Microbiome are provided here courtesy of **BMC**